

Exo-TP2 comptant pour 8%

À remettre le Lundi 18 avril 2016 à 12h00 (MIDI)

Pr. B. Chaib-draa

Question 1 [6pts] : On voudrait représenter à l'aide d'un réseau Bayésien le système de carburant (Fuel) de véhicule à l'aide 3 variables : B pour batterie ; F pour Fuel et G pour jauge. Dans ce contexte, $B = 0$ signifie que la batterie est à plat ; et $B = 1$ qu'elle est chargée. De même $F = 1$ signifie que le réservoir est plein et que $F = 0$ qu'il est vide. Finalement, $G = 1$ signifie plein de carburant et $G = 0$, pas de carburant. Les probabilités sont comme suit :

$$p(B = 1) = 0.9 ; p(F = 1) = 0.9$$

$$p(G = 1 | B = 1, F = 1) = 0.8$$

$$p(G = 1 | B = 1, F = 0) = 0.2$$

$$p(G = 1 | B = 0, F = 1) = 0.2$$

$$p(G = 1 | B = 0, F = 0) = 0.1$$

- a) Donner le réseau bayésien liant B, F, G
- b) Supposons qu'on observe la variable G et qu'on découvre que $G = 0$; quelle est la probabilité pour que le réservoir soit vide ? Commentez cette probabilité par rapport à $p(F = 0)$.
- c) Supposons maintenant qu'on observe à la fois G et B et qu'ils sont tous les deux à 0 ; dites quelle est la probabilité pour que le réservoir soit vide. Commentez cette probabilité par rapport à la probabilité trouvée en b).
- d) Supposons qu'au lieu de lire directement G , c'est le conducteur (D pour Driver) qui se charge de nous dire si G reflète le plein ($D = 1$) ou il reflète le vide ($D = 0$). Donnez alors le nouveau réseau bayésien liant B, F, G, D
- e) En réalité, notre conducteur est peu fiable puisque les probabilités donnent :

$$p(D = 1 | G = 1) = 0.9$$

$$p(D = 0 | G = 0) = 0.9$$

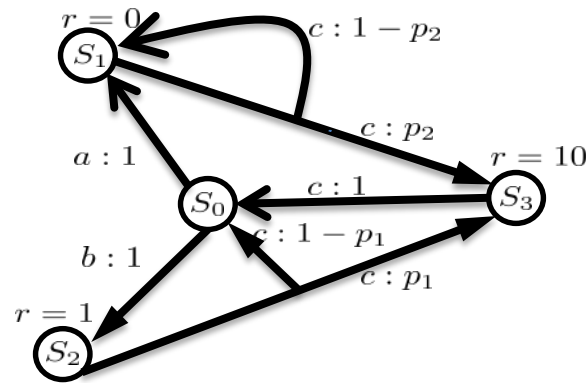
Supposons que le Driver nous disent que $D = 0$; évaluez alors la probabilité pour que le réservoir soit vide sachant uniquement cette observation.

- f) En plus de $D = 0$ on a aussi $B = 0$; évaluez la probabilité pour que le réservoir soit vide et notez que cette probabilité est plus faible que la probabilité déterminée en e). Quelle est l'intuition derrière ça ?

Question 2 [7pts]: Les figures ci-dessous représentent un processus décisionnel de Markov (PDM) où l'effet des actions est stochastique. Chaque arc est associé à une distribution de probabilité de transiter par cet arc en effectuant une action stochastique. La notation du genre $c : 1 - p_2$ signifie que l'action c est effectuée avec une probabilité de $1 - p_2$. La récompense r est de 10 pour l'état S_3 de 1 pour l'état S_2 et il est de 0 ailleurs.

- a) Donner toutes les politiques possibles de ce PDM;
- b) Donner les équations représentant les fonctions de valeurs optimales pour chacun des états : $V^*(S_0)$, $V^*(S_1)$, $V^*(S_2)$, $V^*(S_3)$

- c) Soient $p_1 = p_2 = 0.3$ et $\gamma = 0.9$ calculez V^* et π^* pour tous les états. Pour l'itération de valeurs, vous pouvez adopter l'erreur $\varepsilon = 10^{-3}$ entre 2 valeurs successives ; erreur qui vous permet d'arrêter les calculs.



Question 3 [6pts] : Étant donné les données d'entraînement ci-dessus, concernant la décision de jouer au Tennis ou pas, prédire alors le nouveau exemple, utilisant le classificateur Bayésien naïf : Ciel = Nuageux ; Température = Froide ; Humidité = Normal; Vent = Faible.

Journée	Ciel	Température	Humidité	Vent	JouerTennis
J1	Ensoleillé	Chaude	Élevée	Faible	Non
J2	Ensoleillé	Chaude	Élevée	Fort	Non
J3	Nuageux	Chaude	Élevée	Faible	Oui
J4	Pluvieux	Tempérée	Élevée	Faible	Oui
J5	Pluvieux	Froide	Normal	Faible	Oui
J6	Pluvieux	Froide	Normal	Fort	Non
J7	Nuageux	Froide	Normal	Fort	Oui
J8	Ensoleillé	Tempérée	Élevée	Faible	Non
J9	Ensoleillé	Froide	Normal	Faible	Oui
J10	Pluvieux	Tempérée	Normal	Faible	Oui
J11	Ensoleillé	Tempérée	Normal	Fort	Oui
J12	Nuageux	Tempérée	Élevée	Fort	Oui
J13	Nuageux	Chaude	Normal	Faible	Oui
J14	Pluvieux	Tempérée	Élevée	Fort	Non

Question 4 [5pts]: Étant donnée les mêmes données d'entraînement que précédemment (Question 3), prédisez alors la classe du nouvel exemple en utilisant les K plus proches suivants (avec $K = 4$) : Ciel = Nuageux ; Température = Froide ; Humidité = Normal; Vent = Faible. Pour une mesure de similarité simple utilisez :

$$\text{Similarity}(A, B) = \frac{\sum_{i=1}^4 w_i \cdot \partial(a_i, b_i)}{4}$$

où $\partial(a_i, b_i)$ est égal à 1 si a_i est égale à b_i et 0, autrement. Bien entendu, a_i et b_i reflètent soit Ciel, Température, Humidité ou Vent. Les poids w_i sont tous égaux à 1 sauf pour Température auxquels cas ils sont égaux à 2.

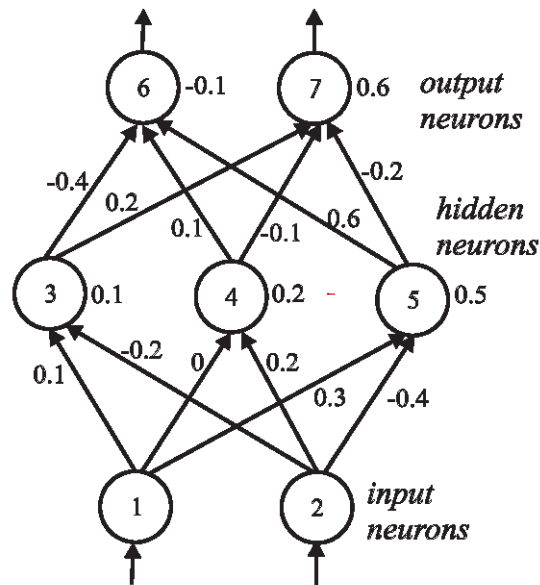
Question 5 [8pts] : Avec les mêmes données d'entraînement ci-dessous, construisez un arbre de décision qui prédit la classe du nouvel exemple : age ≤ 30 , income = medium, student = yes, crédit-rating = fair.

RID	age	income	student	credit_rating	Class: buys_computer
1	≤ 30	high	no	fair	no
2	≤ 30	high	no	excellent	no
3	31 ... 40	high	no	fair	yes
4	> 40	medium	no	fair	yes
5	> 40	low	yes	fair	yes
6	> 40	low	yes	excellent	no
7	31 ... 40	low	yes	excellent	yes
8	≤ 30	medium	no	fair	no
9	≤ 30	low	yes	fair	yes
10	> 40	medium	yes	fair	yes
11	≤ 30	medium	yes	excellent	yes
12	31 ... 40	medium	no	excellent	yes
13	31 ... 40	high	yes	fair	yes
14	> 40	medium	no	excellent	no

Question 6 [8pts]: Soient les données d'entraînement de la table ci-dessous, prédisez en utilisant le classificateur bayésien naïf, la classe du nouveau exemple : windy = false, humidity = 52, temperature = 65, outlook = overcast. **Si les probabilités sont nulles, utilisez Laplace Smoothing.**

outlook	temperature	humidity	windy	play
sunny	85	85	false	no
sunny	80	90	true	no
overcast	83	86	false	yes
rainy	70	96	false	yes
rainy	68	80	false	yes
rainy	65	70	true	no
overcast	64	65	true	yes
sunny	72	95	false	no
sunny	69	70	false	yes
rainy	75	80	false	yes
sunny	75	70	true	yes
overcast	72	90	true	yes
overcast	81	75	false	yes
rainy	71	91	true	no

Question 7 [8pts] : Étant donné le réseau de neurones initialisé avec les poids comme l'indique la figure ci-dessous (où à côté de chaque nœud on a mis le biais b_i et la fonction associée à chacune des sortie est sigmoïde $S(t) = \frac{1}{1 + e^{-t}}$), expliquez alors l'architecture d'un tel réseau sachant que l'on désire distinguer entre Voiture et Camion. Un exemple d'entrainement est à notre disposition, il est comme suit : $T_1\{0.7, 0.15, \text{voiture}\}$ et $T_2\{0.25, 0.4, \text{camion}\}$



Supposons que le taux d'apprentissage α est de 0.15 et que les poids sont indiqués comme dans la figure précédente. Faites alors la propagation en avant des signaux en utilisant T_1 comme entrée ; calculer l'erreur via la propagation arrière. Montrer les calculs et les changements au niveau des poids. À noter que les valeurs juxtaposant les nœuds représentent les **biais** du réseau ; tandis que les valeurs juxtaposant les flèches représentent les **poids** du réseau.